On-road and online:

DATA PRIVACY FOR CONNECTED VEHICLES

Colin McCormick[1]

December 2019

DATA Catalyst

INTRODUCTION

Today's U.S. mass-market cars are increasingly "smart", with on-board wireless data connectivity that provides a range of services. Many of these "connected cars" also carry a growing number of external cameras that record video and images of their surroundings. This inevitably includes pedestrians, cyclists, other drivers, and the license plates of nearby vehicles. If these cars were programmed to aggregate data in a central database about the time and location that individuals and license plates were observed, they would collectively form a constant, rolling mass surveillance network with potentially millions of participating vehicles. This could also take the form of a real-time surveillance system for a "hotlist" of wanted targets through the use of facial recognition or license plate recognition software.

In the absence of federal legislation establishing rights and responsibilities around the ownership and privacy of connected car data, the legal prohibitions against this type of use are unclear. Particularly in the context of the rapid growth of automated license plate readers (ALPRs) and their demonstrated commercial value, auto manufacturers may contemplate monetizing their ability to generate this type of data in the future. Privacy-infringing use of data from connected cars may also come about through cyber intrusions into vehicle software that implement surveillance code but do not affect vehicle operations, making them harder to detect.

This situation has arisen partly because existing data ownership and privacy frameworks for connected cars, both formal and informal, only recognize vehicle owners/operators as a stakeholder. The privacy interests of the general public are not considered, despite the emerging technological capability of connected cars to collect and aggregate relevant data. This situation could be rectified by several forms of Congressional action; states and the automotive industry also have a variety of possible actions they could take.

This paper provides background on the potential for the use of connected cars as surveillance platforms and current thinking on data ownership and privacy for connected cars. It examines several scenarios in detail and concludes with a discussion of potential actions by policymakers and industry to respond.

THE RISE OF CONNECTED CARS

Mass-market cars increasingly incorporate digital technologies designed to make them safer, more efficient, and more entertaining for consumers. Many of these features rely on capturing and using enormous amounts of data, including images, video, audio and geolocation data, from both inside and outside the vehicle. This trend is effectively turning modern vehicles into "rolling computers" that have many aspects in common with mobile digital devices such as smartphones.

Because many of these in-vehicle digital technologies transfer data over wireless communication networks while in motion, these vehicles are often called "connected cars".[2] Approximately 50 million connected cars are on U.S. roads in 2019 (roughly a quarter of the total),[3] with global sales forecast to rise to 72.5 million in 2023.[4] A review by Consumer Reports of 2018 model-year cars found that 32 of 44 automotive brands offered some form of wireless connectivity.[5] The U.S. General Accounting Office (GAO) examined data privacy policies related to connected cars in 2017, and found that 13 of 16 selected automakers offered vehicles with some form of wireless connectivity.[6] It seems clear that the fraction of cars on the road that are network-connected will continue to rise.

Vehicle connectivity is used to provide a range of safety-related services such as automatic crash notification and roadside assistance. Manufacturers also use it to collect vehicle diagnostic data, for purposes that include repair of individual vehicles and issuing vehicle recalls. Connectivity can also support the use of advanced driver assistance systems (ADAS) such as lane-departure warning, although it is not strictly necessary for these functions. In parallel to these safety-related use cases, automakers are rapidly expanding infotainment and in-vehicle marketing and sales applications.[7] The interest from automakers in connectivity is underpinned by the fact that there is substantial market value at stake: by one estimate, "car data" generated in the course of vehicle operation, when monetized, could have a global revenue measured in the hundreds of billions of dollars.[8]

Broadly, connectivity is categorized as either "vehicle-to-infrastructure" (V2I), in which data transfer occurs between the vehicle and stationary systems such as smart traffic lights, or "vehicle-to-vehicle" (V2V), in which data transfer occurs between individual vehicles. Both forms of connectivity are being explored for the wide range of potential benefits they can bring.[9]

CONNECTED CARS AND DATA OWNERSHIP

The growth of connected car applications has led to many unresolved questions about the ownership of car-generated data. Ownership of the data from event data recorders (EDRs) which record technical information before, during and after a crash, was unambiguously assigned to the owner or lessee of a vehicle, and not to the vehicle's manufacturer, by the Driver Privacy Act of 2015.[10] However, the ownership of other data, including geolocation or trip information, infotainment and e-commerce transactions, vehicle diagnostic data, and driving patterns and driver behavior, is not currently addressed by federal legislation.[11]

At the state level, 29 states and the District of Columbia have passed some form of regulation concerning the development of autonomous vehicles, which can be considered a specialized and advanced version of connected cars.[12] However, almost none of these address the issue of data ownership. Diagnostic information is a special case; because of the use of on-board diagnostic (OBD) systems, repair technicians often need access to these data in order to provide maintenance to a vehicle. Control of access to OBD data became a contentious issue earlier than other connected car data, and "right to repair" legislation in Massachusetts led to a 2014 industry agreement to allow open access to vehicle data related to repairs.[13] While notable as an early example of challenges around ownership of connected car data, this and related legislation, as well as relevant industry agreements, have no bearing beyond diagnostic data.

The issue of how "open" car-generated data should be, and the appropriate ownership and governance models for that data, remains a subject of active policy debate both in the U.S.[14] and Europe.[15] It is also the subject of substantial confusion in practice; of 13 major carmakers interviewed by GAO in 2017, seven stated that the ownership of car-generated data is legally unclear or they had no current position on the issue, three stated that the vehicle owner owns the data but the manufacturer has a license to use it, two stated that the manufacturer owns the data, and one stated that the manufacturer owns anonymized data while the vehicle owner owns personally identifiable data.[16]

CONNECTED CARS AND DATA PRIVACY

The issue of rights and responsibilities for privacy of car-generated data is similarly unresolved at the federal level. Although a number of bills have been introduced in Congress (notably the Security and Privacy in Your Car ("SPY Car") Act of 2019) no enacted legislation comprehensively addresses connected car data privacy or establishes a legal framework for the topic.[17]

One reason this is particularly troubling is that car-generated data may be more expansive and represent a greater potential violation of privacy than data collected by mobile phones: in addition to data from the vehicle's own systems such as geolocation and driver behavior, many connected cars attempt to connect to, and read data from, drivers' smart phones, thus gaining access to contact lists, call records, etc.[18] A 2017 proposal by the U.S. National Highway Traffic Safety Administration (NHTSA) to mandate V2V communications in the interests of safety raised significant concerns about enabling privacy-violating tracking of vehicles and was subsequently withdrawn.[19]

The contrast with data privacy law for mobile phones is striking; that topic has risen to the level of the Supreme Court, which ruled in Carpenter v. United States in 2018 that law enforcement generally needs a warrant to obtain location information derived from mobile phones.[20] In Byrd v. United States the Supreme Court ruled that drivers not listed on rental-car contracts still have reasonable expectations of privacy while driving. An amicus brief in the case filed by the Electronic Privacy Information Center (EPIC) and a number of legal and technical experts highlighted the fact that "the warrantless search of a modern vehicle implicates far more privacy interests than the physical search of a '66 Buick LeSabre" because of the enormous amount of data stored in connected vehicles.[21] However, the ruling did not address the broader issue of vehicle data privacy.

Even the jurisdiction of federal agencies over the topic of connected car data privacy is unclear. NHTSA has asserted some degree of regulatory authority, issuing guidance in 2016 for vehicle cybersecurity best practices that included considerations of data privacy.[22] However, the Federal Trade Commission (FTC) has primary jurisdiction in matters of consumer data protection. NHTSA and FTC co-hosted a workshop in 2017 to address connected car data issues generally, including data privacy, but have not taken any joint action since that time.[23]

In 2014, the Alliance of Automobile Manufacturers (the "Auto Alliance"), representing a range of major auto manufacturers, established a set of consumer data privacy principles related to data collected by in-car technologies, to which a large number of automakers have publicly committed to "meet or exceed".[24] A non-binding "Resolution on data protection in automated and connected vehicles" was adopted in 2017 by the International Conference of Data Protection and Privacy Commissioners (ICDPPC), calling on a range of stakeholders including standards bodies, governments, and manufacturers to address issues of data privacy.[25]

WHOSE PRIVACY SHOULD BE PROTECTED?

Throughout the debates surrounding connected car data ownership and data privacy responsibilities, there is a consistent assumption by participants that there are only two categories of stakeholders who must be considered: the owner, operator and passengers of the vehicle on the one hand, and the auto manufacturer on the other. This is reflected in the Driver Privacy Act of 2015 (which assigned all EDR rights to the owner), the Auto Alliance consumer data principles (which apply only to the data privacy rights of vehicle owner/operators), the data-protection resolution adopted by the ICDPPC (which urges consideration of the rights of users of the vehicle), and the views on the topic expressed by auto manufacturers to GAO.

However, there is one more stakeholder whose data privacy rights are at issue in the context of modern connected cars: the public at large. Connected vehicles increasingly carry high-resolution cameras (often more than one) and record video and images that capture the faces of pedestrians, cyclists and other drivers in their vicinity, as well as license plates. Modern facial- and object-recognition algorithms are able to use these images to identify specific individuals and read specific license plates. This capability, combined with the vast and growing number of camera-equipped connected cars with wireless network capability, could give rise to an enormous system of mass surveillance. The subjects of this surveillance would not be the owners, operators and passengers of connected vehicles, but rather the individuals who are in the vicinity of those vehicles during their normal operation.

Early academic work on the topic of data privacy for autonomous vehicles – whose sensor systems are far more advanced than the cameras used by conventional connected cars – examined the question of whether networked vehicles could conduct mass surveillance of their passengers.[26] Later work expanded on this topic, noting that networked autonomous vehicles could also potentially identify individuals in their vicinity to create a form of mass surveillance, and examined public attitudes towards this situation.[27] However, these efforts focused on the capabilities of highly advanced autonomous vehicles, which are still largely in the experimental phase[28] with approximately 1,400 operating as test vehicles in the U.S.[29] Because connected cars in the U.S. number in the tens of millions, they pose a much greater and more immediate concern as the potential basis of a mobile mass surveillance system.

CAMERAS ON CONNECTED VEHICLES

To better understand this issue, it is important to note that many vehicles now contain multiple external and internal cameras.[30] A prominent example of this is rear-view ("backup") cameras, which were federally mandated in all new US light-duty vehicles as of 2018 because of their ability to help drivers avoid collisions or other accidents when reversing.[31] Backup cameras are generally used in a "direct" mode such that the driver directly watches the video image while driving with no automated processing, although virtual line markings are often digitally added to the image based on the steering wheel angle and automated detection of the range to a nearby vehicle, such as during parallel parking. The field of view of these cameras varies by manufacturer, but the license plate of trailing vehicles is often visible, as well as individuals who are in the vicinity of the rear of the car and potentially the driver and passengers of a trailing vehicle.

Forward-facing cameras, usually mounted by the rear-view mirror or on the dashboard, are not mandated but are increasingly common. They are one of the primary sensors that supports ADAS, the suite of safety-related services that includes lane-departure warning (a notification to the driver if the vehicle is moving out of a marked lane without the turn signal being activated) and adaptive cruise control (automatic detection of the distance to a leading vehicle with corresponding acceleration/braking to maintain fixed separation).[32] The images and video from these cameras are not used in a direct mode – the driver does not view them directly while driving – but are instead processed by advanced computer vision algorithms to extract information such as the range and speed of a leading vehicle. This information is then used by operational algorithms to either provide driver notification (as in lane departure) or actively control the vehicle (as in adaptive cruise control). These cameras usually have a clear field of view of objects in front of the vehicle, including license plates and individuals. Some after-market "dashcam" systems are also used to record video footage for liability and insurance purposes in the event of a crash, but these are not necessarily networked and are usually not integrated into the connected car's main systems.

Some vehicle models include additional external cameras beyond forward-facing and backup cameras to provide a "360-degree" view of the surrounding area. These are intended to provide information about the traffic environment to the side as well as ahead and behind; for example, several Tesla models use up to eight cameras for a variety of purposes.[33] The automotive industry has generally embraced the use of cameras to provide this broader view of the vehicle's surroundings. NHTSA has recently announced plans to test rear- and side-mounted cameras as complete replacements for conventional physical mirrors, following a 2014 petition from the automotive industry.[34] This and other uses of side-mounted external cameras would likely blend both direct (driver-viewed) and indirect (algorithmically processed) modes of use. Overall, the clear trend is for vehicles to incorporate a growing number of external cameras that capture imagery and video of a large fraction of the vehicle's surroundings, with a significant fraction of this imagery being analyzed by computer algorithms to extract information.

EXTERNAL CAMERAS AND NON-OWNER PRIVACY: DATA RETENTION

The rapid increase in vehicle-mounted external cameras means that many connected cars now have an enormous capacity for recording images and videos from their surroundings, both as they drive and while parked. Since this inevitably includes faces of pedestrians, cyclists and other drivers, as well as license plates, it would be relatively straightforward for algorithms to identify individuals and/or read license plates, particularly for imagery that is already processed by algorithms to be used in in "indirect" mode. When these data are time- and location-stamped, they become a record of individual activity, or can be linked to individuals in the case of license plates.

Below, I present two fundamental scenarios by which data recorded by connected-vehicle-mounted cameras could pose a threat to privacy as part of a mass-surveillance system. The first is based on massive data retention. The second is real-time identification.

Threats from mounted cameras: Scenario 1 – massive data retention

If time- and location-stamped images or video footage are transmitted by the vehicle to be stored in databases maintained by the manufacturer (via the vehicle's wireless connection) this creates a lasting record that can be interrogated to retroactively locate specific individuals. By combining camera data from a very large number of vehicles, which would be the case for datasets accumulated by manufacturers ("fleet aggregation"), this could become a detailed and extensive record of individual movements over time, far beyond a single instance of photographing an individual at one time and location. These same observations apply in the case of license plates; time- and location-stamped images that include a clear view of license plates effectively serve as a record of the movements of those specific vehicles. An important aspect to note is that this scenario could occur even without advanced object- and facial-recognition algorithms operating on board the connected cars. Simply recording and retaining raw video and image data in a central database, and searching it later with these algorithms, would be inefficient but technologically possible.

While U.S. law generally allows capturing images of individuals in public,[35] the law also establishes a "reasonable expectation of privacy" for daily movements.[36] A system of mass surveillance that used images and video from connected car external cameras across a fleet of vehicles could aggregate data on individual and vehicle movements across an enormous range of places and times. Using the taxonomy of privacy developed by Daniel Solove, this scenario would clearly constitute (at a minimum) surveillance, aggregation and identification.[37] Analyzing and searching this aggregated data could lead directly to unreasonable, warrantless searches that would violate the Fourth Amendment.[38]

This general issue is not new: as early as 2008, Google responded to privacy concerns about faces recorded in images taken from vehicle-mounted cameras for Street View and implemented technology to blur them out (although this did not catch all such instances).[39] [40] However, the privacy implications of Street View are mitigated by the infrequent collection of data (Street View images are updated only rarely) and the very small size of the vehicle fleet used, which is somewhat analogous to the situation with fully autonomous vehicles.[41]

The technique of fleet aggregation is also not new. Many automakers and related companies are developing datasets based on aggregated data collected by fleets of autonomous vehicles being tested on public roads, including camera images and video. Notably, Alphabet's Waymo has collected data from over ten million miles of on-road driving by test vehicles in 25 U.S. cities, including high-resolution video from multiple cameras.[42] A portion of this data has been released for research purposes; while faces and license plates appear to have been obscured in the released data, it is unclear what fidelity of data is saved in the company's internal datasets.[43] Tesla, which is developing a form of autonomous driving technology based on data from tens of thousands of ordinary (i.e. non-test) vehicles, has accumulated over one billion miles of on-road driving data, which appears to include a similarly large amount of high-resolution video; the treatment of personally identifiable information in this dataset is also unclear.[44]

A related concern is the fate of data recorded by external cameras when vehicles are sold or junked. Many connected cars have substantial data-storage capabilities and permanently save a variety of information by default.[45] This data can become exposed when the car is sent to a junkyard or otherwise disposed or sold; this was demonstrated in 2019 when video footage was recovered from a Tesla vehicle that was purchased from a junkyard.[46] The U.S. Federal Trade Commission (FTC) has issued guidelines for clearing personal data from vehicles before selling them, but they address only owner-related data, not any data (such as video) that potentially relates to other individuals.[47]

Threats from mounted cameras: Scenario 2 – real-time identification

The second scenario for a connected-car-enabled mass surveillance system concerns real-time identification. This would be enabled by the increasing capability of computer vision algorithms to detect and analyze human faces and license plates in real time.[48] In addition to detection (in which the algorithm flags the presence or absence of the target, and its location within the image frame) related algorithms are able to identify the specific individual whose face was captured in the frame (facial recognition) and read the letters and numbers of the license plate captured in the frame (license plate recognition).

ADAS systems operating in connected cars currently incorporate sophisticated computer vision algorithms and embedded processors, which are used for object detection and other purposes.[49] Fleets of these vehicles could be programmed to search for a specific individual or license plate in real time and notify the manufacturer or other entity of a positive identification. This would create a potential privacy violation in real-time, as opposed to the retroactive searching of datasets to establish the previous movements of an individual or vehicle. The decision in *Carpenter* held that obtaining historical cell-site location information without a warrant was a Fourth Amendment search requiring a warrant, while leaving unaddressed the issue of obtaining real-time cell-site location data.[50] A similar dichotomy holds in this case: aggregating historical location information through cameras on-board connected cars is distinct from (although not mutually exclusive with) the use of ADAS and other external cameras to identify the current location of specific individuals.

A key aspect to the privacy concern articulated here is the collection of data derived from external cameras from a large number of connected vehicles, essentially a "crowd-sourced" or "fleet-sourced" technique. Several versions of this scenario (not involving privacy-related data) are already taking place for the purposes of building and updating ultra-high-resolution road map datasets, including Intel subsidiary Mobileye, which uses ADAS cameras to gather real-time information about roads being driven by connected cars.[51]

THE SPECIAL CASE OF EXTERNAL CAMERAS AND LICENSE PLATES

The issue of potential privacy concerns with external cameras is particularly important for the case of license plates, for a variety of reasons. From a technical point of view, implementing computer vision algorithms for identifying license plates is an easier computational problem than human facial recognition, because license plates use relatively standardized sizes and fonts, and have standardized placement on the vehicle.[52] While potential concerns surrounding the ability of external cameras to perform human facial recognition should not be neglected, it appears that adding software to ADAS systems or other external cameras to read license plates would be relatively simpler.

Additionally, the use of automated license plate recognition has grown dramatically in recent years, particularly by law enforcement. Commercially available devices known as automated license plate readers (ALPRs) use cameras mounted on police vehicles, light poles, or other locations to take pictures of license plates on nearby vehicles (parked or in motion), recording the time and location of each one. They then apply computer vision algorithms to automatically read the letters and numbers of the plate and compare them with "hotlists" maintained by law enforcement to try to identify vehicles that are stolen or otherwise associated with criminal activity.[53] ALPRs are able to read over 1,000 plates per minute (although actual use conditions usually mean far fewer license plates pass the camera's field of view during this time) and at vehicle speeds as high as 120 to 160 mph.[54] Costs of ALPRs have fallen from over $10,000 in 2011[55] to hundreds of dollars today, due to a growing ability to use off-the-shelf cameras and lower software licensing fees.[56] They have demonstrated the ability to enhance police effectiveness for a number of vehicle-related crimes, although a systematic review of their impact is lacking[57] and the use cases continue to expand.[58]

The use of ALPRs has led to extensive concerns about violation of privacy, particularly due to the ongoing growth of databases that combine data about the time and location of all vehicles detected by multiple ALPRs.[59] [60] The largest known license plate database, containing billions of records of vehicle locations, is maintained by a private company and used under contract by the U.S. Immigrations and Customs Enforcement (ICE) agency to identify people who may be in the United States illegally.[61] There is no federal policy regarding the use of ALPRs in place, state policies governing their use vary greatly,[62] and recent court rulings have not provided clarity.[63] ALPRs also raise major cybersecurity concerns. In 2015, researchers associated with the Electronic Frontier Foundation found a large number of online law-enforcement ALPR systems whose data were partly or entirely unprotected;[64] many of these systems were found to still be vulnerable four years later.[65]

Within this context, it seems reasonable to ask whether law enforcement might request access to an auto manufacturer's ADAS cameras on connected cars for the purpose of automatic license plate recognition, expanding the existing network of ALPRs to include a vastly larger number of "fleet-sourced", vehicle-mounted ones. Under this scenario, law enforcement could provide a "hotlist" of license plates to auto manufacturers and request that ADAS cameras on connected cars be programmed to watch for them and report any sightings in real time. Alternatively (although not exclusively) law enforcement could request that auto manufacturers implement automatic license plate recognition through ADAS cameras and store the resulting data to be searched in the future. Because the technical capability appears to be in place in a growing number of vehicles, and there is clearly a perceived utility of license plate recognition by law enforcement, such a scenario is not inconceivable at some point in the future.

SILENT SPYWARE: CONNECTED CARS AND CYBERSECURITY

The scenarios described above envision that data aggregation and/or on-board image analysis using specialized algorithms would be implemented by auto manufacturers. However, it is possible that software to aggregate and/or analyze images, and transmit the resulting information to a central database, could be implemented by a third party by attacking a vehicle's software systems, potentially by compromising an over-the-air (OTA) software update.[66] Under this scenario, connected cars could implement "fleet aggregation" surveillance without the knowledge of the manufacturer or the owner/operator. The third party implementing this attack could be interested in the commercial value of the resulting data or could be a state actor interested in surveillance in particular locations such as outside military or other national-security facilities. A "silent spyware" version of this scenario could also include remotely inserted mass surveillance code in a large number of connected cars that is not currently active but could be activated at some point in the future to operate for a brief period of time before being detected.

Mitigating against this scenario is particularly difficult. Because of the many digital services they provide, connected cars run software that includes millions of lines of code.[67] Software of this complexity creates a major challenge for auto manufacturers to ensure cybersecurity. An additional factor in this challenge is the fact that software components are distributed over many sub-systems (electronic control units, or ECUs) that provide services ranging from ignition timing to power window control to braking. These are manufactured by a wide variety of suppliers, many of whom are outside the US. The cybersecurity threat for connected cars was vividly demonstrated in 2015 when security researchers remotely took control of a Jeep Cherokee and ran it off the road, via the vehicle's connected entertainment system.[68] Other examples include similar attacks on Ford and Toyota vehicles,[69] as well as multiple attacks on Tesla vehicles.[70]

Many analysts have since examined the broader implications and overall threat posed by vehicle connectivity.[71] The automotive industry has responded to these concerns through efforts such as the Automotive Information Sharing and Analysis Center (Auto-ISAC), but the issue of ensuring cybersecurity for modern vehicles is far from solved.[72] Notably, it will only become more difficult with the advent of fully autonomous vehicles, where even the basic frameworks for assessing cyber vulnerabilities of vehicles are incomplete.[73]

POLICY AND INDUSTRY RESPONSES

There is a wide range of possible responses to the scenarios outlined above, from both policymakers and the automotive industry. At the federal level, Congress is considering a variety of options for a federal data privacy law.[74] This topic is large and politically challenging, but it may provide an opportunity to specifically address privacy and ownership for connected car data. Should Congress choose to address this issue, it will be important to consider data rights and responsibilities for three categories of stakeholders: the owner/operator/passengers of connected cars, auto manufacturers, and the general public in the vicinity of such cars whose personally identifiable information is or could be collected through the operation of the vehicle's ADAS or other sensor systems, including facial recognition and license plate recognition.

More broadly, Congress could choose to establish the underlying principle that the increased safety benefits of external cameras on vehicles are directly linked to increased data-privacy responsibilities on the part of the automaker. Doing so would require addressing how tradeoffs between increased safety and increased risks to data privacy should be weighed by federal agencies, similar to the situation encountered by NHTSA in 2017 in contemplating a V2V mandate.

Congress could instead choose to address the issue with narrower legislation, including various bills that have already been introduced. Similar considerations as the above apply in this case, although within a narrow legislative context there may be more room to establish rights and responsibilities for various categories of connected car data, such as distinguishing among diagnostics, geolocation, and infotainment data. However, a more robust legislative strategy may be to direct NHTSA and FTC to establish and regularly update definitions around these categories of connected car data as technology evolves, with associated rights and responsibilities established through rulemaking. This could include explicit consideration of the privacy rights of the public as it pertains to connected car external cameras. Congress could additionally choose to clarify or change the responsibilities of NHTSA and FTC on the topic of connected car data ownership and privacy, but in so doing would need to articulate clear delineations for connected car data as distinct from other consumer data.

The special case of the scenario in which external cameras on connected cars could be used as "mobile license plate readers" could be addressed through legislation related to law enforcement practices. It could directly limit the market for license plate data by directing ICE to restrict or end its use of this type of data or establish strong controls over data sources and providers. More broadly, Congress could pass legislation establishing a federal policy on the use of automated license plate readers by law enforcement and explicitly restrict civilian vehicle external cameras from being operated in this way. Congress could also consider policies based on federal funding, investigation of law enforcement practices, and other avenues.[75]

On the issue of "silent spyware", Congress could choose to enhance funding for vehicle cybersecurity efforts, and direct federal agencies such as the Department of Transportation (DOT) and the National Institute of Standards and Technology (NIST) to expand their vehicle cybersecurity work around detection of software intrusions that result in surveillance rather than vehicle operational changes.

At the state level, the California Consumer Privacy Act (CCPA) provides the strongest data-protection framework currently in place in the US (the law goes into effect in 2020) and is likely to have a major impact on data-protection practices across many sectors. It will be directly relevant to connected car data, and it may be interpreted by auto manufacturers as establishing a requirement to obtain consent from vehicle owners to use certain forms of data collected by the vehicle. The California legislature or the attorney general could stipulate that under the law the data privacy of the general public is included within the responsibilities of auto manufacturers as it pertains to connected cars. California could also pursue this issue through convening workshops among auto manufacturers, privacy advocates, and automobile safety organizations, possibly working with NHTSA and/or FTC. Other states could develop independent policy agendas on the topic of connected car data privacy, but experience has shown that many states prefer to act in concert on regulatory topics of this character and thus may prefer to follow California's lead.

Auto manufacturers could establish leadership on this topic by expanding consumer data privacy principles to include commitments to individuals in the vicinity of operating connected vehicles whose personally identifiable information (including facial recognition and license plate recognition) could be collected through normal vehicle operations. Simply acknowledging the privacy implications to the general public from the operation of cameras on connected cars could be a helpful contribution to the policy debate, by building public awareness of the issue and highlighting the need to clarify the legal situation around connected car data privacy. This could also catalyze further dialogue about the tradeoffs between vehicle safety and the privacy of the general public. Auto manufacturers could also offer public commitments that their vehicles will not be used as platforms for broad surveillance, either in real time or retroactively through aggregation of data; this commitment could include assurances against commercial use of this form of surveillance as well as reasonable protections against cyber-intrusions for surveillance purposes. Further, auto manufacturers could implement and publicize responsible data-retention policies, in which collected image and video data are only stored in privacy-protected fashion (which is likely to include, at a minimum, face and license-plate blurring similar to Street View) and retained for as short a time as possible.

Ultimately, the technical complexity of this topic and the large number of stakeholders involved means that cooperative solutions will likely be the most effective. Auto manufacturers, regulators and privacy advocates could hold periodic convenings to explore the appropriate balance between improving automotive safety and services through the increased use of external cameras and protecting reasonable standards of privacy for individuals including pedestrians, bicyclists and other drivers. The federal government could provide funding support and explicitly link the outcomes of discussions in this forum to informing further federal legislation on connected car data privacy.

CONCLUSION

Connected cars are increasingly common on American roads and provide many digital services to their occupants that enhance the safety and enjoyability of driving. Because these vehicles often use external cameras, they could potentially become platforms for mass surveillance of individuals and vehicles in their vicinity. The impact of this for a small number of vehicles would be negligible but given the tens of millions of connected cars currently in use (and projections for continued growth) the possible implications of a truly mass surveillance system based on fleets of ordinary vehicles with network connectivity become far more significant.

Because the technological components of this type of mass surveillance system appear to already be in place, it is important to consider the scenarios under which it might be implemented. Auto manufacturers, in the absence of clear regulatory requirements about the data privacy of the general public in the vicinity of connected cars, could choose to install software that would use videos and images from vehicle external cameras for facial recognition and/or license plate recognition. This information would likely have substantial commercial value, and the potential scale of such a system would be enormous compared with other surveillance systems in place. Alternatively, third parties could exploit cyber vulnerabilities in connected car software to install "silent spyware" that would use connected car external cameras for surveillance without the knowledge of auto manufacturers or vehicle operators.

Congress, the states, and the automotive industry can all take action to avoid these scenarios from occurring. Congress could address connected car data privacy in the context of broader federal data privacy legislation or as a narrower, sector-specific bill; this could include specific attention to the data privacy of the general public in the context of connected cars. At the state level, the California Consumer Privacy Act is the foremost data-protection framework currently in place in the US and will likely impact connected car data privacy. California could explicitly highlight data privacy for the general public as related to connected cars as part of the law's implementation, and other states could choose to adopt this approach or otherwise partner with California. Within the automotive industry, manufacturers of connected cars could extend consumer data rights principles to include the general public and commit to preventing the use of their vehicles for mass surveillance either for commercial reasons or through cyber intrusions.

ENDNOTES

[1] Professor, Science, Technology and International Affairs, Walsh School of Foreign Service, Georgetown University; and Chief Technologist, Valence Strategic LLC.

[2] Elliot, D. et al. "Recent advances in connected and automated vehicles", *J. Traf. Trans. Eng.* (2019). https://www.sciencedirect.com/science/article/pii/S2095756418302289

[3] Data from Statista website (accessed October 7, 2019). https://www.statista.com/outlook/320/109/connected-car/united-states

[4] IHS Markit, The Connected Car (2018). https://ihsmarkit.com/topic/autonomous-connected-car.html

[5] Plungis, J. Who owns the data your car collects? *Consumer Reports* (2018). https://www.consumerreports.org/automotive-technology/who-owns-the-data-your-car-collects/

[6] U.S. Government Accountability Office. Vehicle Data Privacy: Industry and federal efforts underway, but NHTSA needs to define its role, *GAO-17-656* (2017). https://www.gao.gov/assets/690/686284.pdf

[7] Examples of in-vehicle marketing and sales applications include GM's Marketplace app and Honda's Dream Drive app; these and others allow drivers and passengers to buy food, make reservations, and perform a growing list of other online activities.

[8] Monetizing car data: New service business opportunities to create new customer benefits, *McKinsey & Company* (2016). https://www.mckinsey.com/~/media/McKinsey/Industries/Automotive%20and%20Assembly/Our%20Insights/Monetizing%20car%20data/Monetizing-car-data.ashx

[9] Chang, J. et al. Estimated benefits of connected vehicle applications: dynamic mobility applications, AERIS, V2I safety, and road weather management. *FWHA-JPO-15-255, US Department of Transportation* (2015). https://rosap.ntl.bts.gov/view/dot/3569

[10] The Driver Privacy Act of 2015 is a subtitle of the Fixing America's Surface Transportation (FAST) Act of 2015, which became P.L. 114-94. https://www.congress.gov/bill/114th-congress/house-bill/22/text

[11] Zhang, S. Who owns the data generated by your smart car? *Harvard Journal of Law & Technology* (2018). http://jolt.law.harvard.edu/assets/articlePDFs/v32/32HarvJLTech299.pdf

[12] National Conference of State Legislatures, Self-driving vehicles enacted legislation website (accessed November 4, 2019). http://www.ncsl.org/research/transportation/autonomous-vehicles-self-driving-vehicles-enacted-legislation.aspx

[13] Alliance of Automobile Manufacturers, Association of Global Automakers, Automotive Aftermarket Industry Association, and Coalition for Auto Repair Equality, Memorandum of Understanding: R2R Agreement (2014). https://www.nastf.org/files/public/OtherReference/MOU_SIGNED_1_15_14.pdf

[14] Determann, L. and Perens, B. Open Cars. *Berkeley Technology Law Journal* (2017).

[15] Kerber, W. Data governance in connected cars: The problem of access to in-vehicle data. *Journal of Intellectual Property, Information Technology and E-Commerce Law* (2018). http://btlj.org/data/articles2017/vol32/32_2/determann_web.pdf

[16] GAO, op. cit.

[17] McQuinn, A. and Castro, D. A policymaker's guide to connected cars. *ITIF* (2018). http://www2.itif.org/2018-policymakers-guide-connected-cars.pdf

[18] Colbert, C. Privacy under the hood: Towards an international data privacy framework for autonomous vehicles, *We Robot 2018* (2018). https://conferences.law.stanford.edu/werobot/wp-content/uploads/sites/47/2018/02/Privacy-Under-the-Hood-Towards-an-International-Data-Privacy-Framework-for-Autonomous-Vehicles.pdf

[19] Electronic Frontier Foundation, Comments on NPRM: Federal Motor Vehicle Safety Standards; *V2V Communications, Docket No.*

NHTSA-2016-0126 (2017). https://www.eff.org/document/eff-comments-nhtsa-re-v2v-notice-proposed-rulemaking

[20] Carpenter v. United States (2018). https://www.supremecourt.gov/opinions/17pdf/16-402_h315.pdf

[21] Electronic Privacy Information Center, amicus brief, Byrd v. United States (2017). https://epic.org/amicus/fourth-amendment/byrd/Byrd-v-US-EPIC-Amicus-Brief.pdf

[22] NHTSA, Cybersecurity best practices for modern vehicles, Report No. DOT HS 812 333 (2016). https://www.nhtsa.gov/staticfiles/nvs/pdf/812333_CybersecurityForModernVehicles.pdf

[23] Federal Trade Commission, "Connected Cars Workshop: Staff Perspective" (2018). https://www.ftc.gov/system/files/documents/reports/connected-cars-workshop-federal-trade-commission-staff-perspective/staff_perspective_connected_cars_0.pdf

[24] See the Automotive Consumer Privacy Protection Principles website (accessed October 4, 2019). https://autoalliance.org/connected-cars/automotive-privacy/

[25] International Conference of Data Protection and Privacy Commissioners, "Resolution on data protection in automated and connected vehicles" (2017). https://edps.europa.eu/sites/edp/files/publication/resolution-on-data-protection-in-automated-and-connected-vehicles_en_1.pdf

[26] Glancy, D. Privacy in autonomous vehicles. *Santa Clara Law Review* (2012). https://digitalcommons.law.scu.edu/cgi/viewcontent.cgi?article=2728&context=lawreview

[27] Bloom, C. et al. Self-driving cars and data collection: Privacy perceptions of networked autonomous vehicles. *Proceedings of the 13th Symposium on Usable Privacy and Security* (2017). https://www.usenix.org/conference/soups2017/technical-sessions/presentation/bloom

[28] McCormick, C. Self-driving cars find their way in the world. *Physics Today* (2019). https://physicstoday.scitation.org/doi/10.1063/PT.3.4256

[29] See Remarks Prepared for Delivery by U.S. Secretary of Transportation Elaine L. Chao, UBER Elevate Symposium (2019). https://www.transportation.gov/briefing-room/uber-elevate-symposium

[30] See MyCarDoesWhat (accessed October 3, 2019). https://mycardoeswhat.org/

[31] See Federal Motor Vehicle Safety Standards; Rear Visibility (79FR19178), Final Rule. The compliance date for the rule was May 1, 2018. https://www.govinfo.gov/content/pkg/FR-2014-04-07/pdf/2014-07469.pdf

[32] See Automated Vehicles for Safety website, NHTSA (accessed October 7, 2019). https://www.nhtsa.gov/technology-innovation/automated-vehicles-safety

[33] See, e.g. Future of Driving, Tesla (accessed October 3, 2019). https://www.tesla.com/autopilot

[34] Pietsch, B. U.S. to test mirrorless, camera-based systems in autos. *Reuters* (2019). https://www.reuters.com/article/us-usa-auto-mirrorless/u-s-to-test-mirrorless-camera-based-systems-in-autos-idUSKCN1VH2G9

[35] Manning, M.L., "Less than Picture Perfect: The Legal Relationship between Photographers' Rights and Law Enforcement." *Tennessee Law Review* (2010). https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1857623

[36] American Bar Association, Privacy in an Interconnected World, (2019). https://www.americanbar.org/groups/gpsolo/publications/gp_solo/2011/june/privacy_in_an_interconnectedworld/

[37] Solove, D. A taxonomy of privacy. *University of Pennsylvania Law Review* (2006). https://www.law.upenn.edu/journals/lawreview/articles/volume154/issue3/Solove154U.Pa.L.Rev.477(2006).pdf . This taxonomy is a helpful method to organize discussions of privacy violation and reduce them to identifiable categories. These include (a) information collection (surveillance and interrogation), (b) information processing (aggregation, identification, insecurity, secondary use and exclusion), (c) information dissemination (breach of confidentiality, disclosure, exposure, increased accessibility, blackmail, appropriation and distortion), and (d) invasion

(intrusion and decisional interference).

[38] Berman, E. When database queries are Fourth Amendment searches. *University of Minnesota Law Review* (2018). https://minnesotalawreview.org/article/when-database-queries-are-fourth-amendment-searches/

[39] Shankland, S. Google begins blurring faces in Street View. *C|net* (2008). https://www.cnet.com/news/google-begins-blurring-faces-in-street-view/

[40] Frome, A. et al. Large-scale privacy protection in Google Street View. *IEEE Int. Conf. Comp. Vis.* (2009). https://ai.google/research/pubs/pub35481

[41] See Google Street View policy website (accessed October 7, 2019). https://www.google.com/streetview/policy/

[42] Krafcik, J. Where the next ten million miles will take us. *Medium* (2018). https://medium.com/waymo/where-the-next-10-million-miles-will-take-us-de51bebb67d3

[43] Hawkins, A. Waymo is making some of its self-driving car data available for free to researchers. *The Verge* (2019). https://www.theverge.com/2019/8/21/20822755/waymo-self-driving-car-data-set-free-research

[44] Hull, D. Tesla customers rack up 1 billion miles driven on Autopilot. *Bloomberg* (2018). https://www.bloomberg.com/news/articles/2018-11-28/tesla-customers-rack-up-1-billion-miles-driven-on-autopilot

[45] One industry expert estimated in 2015 that connected cars used between 16 GB and 256 GB of on-board memory, approximately the same as a typical laptop. Since that time, the data-storage industry has introduced increasingly larger automotive-grade storage devices, with the first 1 TB drive entering the market in 2018. See, e.g. Harrison, J. Memory use in automotive, *Electronic Products* (2015) https://www.electronicproducts.com/Digital_ICs/Memory/Memory_use_in_automotive.aspx ; and Micron press release (2018) http://investors.micron.com/news-releases/news-release-details/micron-introduces-industrys-first-1tb-automotive-and-industrial

[46] Fazzini, K. and Kolodny, L. Tesla cars keep more data than you think, including this video of a crash that totaled a Model 3. *CNBC* (2019). https://www.cnbc.com/2019/03/29/tesla-model-3-keeps-data-like-crash-videos-location-phone-contacts.html

[47] Tressler, C. Selling your car? Clear your personal data first. *U.S. Federal Trade Commission* (2018). https://www.consumer.ftc.gov/blog/2018/08/selling-your-car-clear-your-personal-data-first

[48] Grother, P. et al. Ongoing facial recognition vendor test (FRVT) Part 2: Identification. *NISTIR 8238* (2018). https://nvlpubs.nist.gov/nistpubs/ir/2018/NIST.IR.8238.pdf

[49] See, e.g., Mody, M. ADAS front camera: Demystifying resolution and frame rate. *EE News* (2016). https://www.eetimes.com/author.asp?section_id=36&doc_id=1329109 ; and Pickering, P. ADAS applications drive advances in image-processing architectures. *Innovation Destination* (2018). https://innovation-destination.com/2018/03/14/adas-applications-drive-advances-in-image-processing-architectures/

[50] Howe, A. Opinion analysis: Court holds that police will generally need a warrant for sustained cell phone location information. *SCOTUS blog* (2018). https://www.scotusblog.com/2018/06/opinion-analysis-court-holds-that-police-will-generally-need-a-warrant-for-cellphone-location-information/

[51] Mobileye RoadBook: high-precision HD maps distributed at scale, *Medium* (2018). https://blog.mapbox.com/mobileye-roadbook-high-precision-hd-maps-distributed-at-scale-9b4e692d29f

[52] Both license plate reader software and facial recognition software are openly available (e.g. OpenALPR, https://github.com/openalpr/openalpr ; and OpenFace https://cmusatyalab.github.io/openface/).

[53] Gierlack, K. et al. License plate readers for law enforcement. *RAND Corporation* (2014). https://www.rand.org/pubs/research_reports/RR467.html

[54] Roberts, D. and Casanova, M. Automated license plate recognition (ALPR) use by law enforcement: Policy and operational guide. *Final grant*

report, US Department of Justice (2012). https://www.ncjrs.gov/pdffiles1/nij/grants/239605.pdf

[55] Phillips, J. Smile! Your car's on camera. *Car and Driver* (2011). https://www.caranddriver.com/features/a15122481/smile-your-cars-on-camera-we-ride-along-to-learn-what-the-cops-know-about-you-feature/

[56] Kaplan, J. License plate readers are creeping into neighborhoods across the country. *Slate* (2019). https://slate.com/technology/2019/07/automatic-license-plate-readers-hoa-police-openalpr.html

[57] Rausch, S. The rise of license plate reader technology. *Security Magazine* (2019). https://www.securitymagazine.com/articles/90188-the-rise-of-lincese-plate-reader-technology

[58] Gierlack, op. cit.

[59] Zmud, J. et al. License plate reader technology: Transportation uses and privacy risks. *Texas A&M University* (2016). https://pdfs.semanticscholar.org/74c1/da6465fabc44e64c714819c7561ed553b1ef.pdf

[60] See, e.g., You are being tracked, *ACLU website* (accessed October 7, 2019). https://www.aclu.org/issues/privacy-technology/location-tracking/you-are-being-tracked

[61] See Friedersdorf, C. An unprecedented threat to privacy. *The Atlantic* (2016). https://www.theatlantic.com/politics/archive/2016/01/vigilant-solutions-surveillance/427047/ ; and Harwell, D. and Romm, T. ICE is tapping into a huge license-plate database, ACLU says, raising new privacy concerns about surveillance. *Washington Post* (2019). https://www.washingtonpost.com/technology/2019/03/13/ice-is-tapping-into-huge-license-plate-database-aclu-says-raising-new-privacy-concerns-about-surveillance/

[62] See Automated license plate readers: state statutes, *National Conference of State Legislatures website* (accessed October 7, 2019). http://www.ncsl.org/research/telecommunications-and-information-technology/state-statutes-regulating-the-use-of-automated-license-plate-readers-alpr-or-alpr-data.aspx

[63] Atiyeh, C. License-plate readers are dealt a blow in Virginia, but privacy is still a rare commodity nationwide. *Car & Driver* (2019). https://www.caranddriver.com/news/a27059427/license-plate-readers-lpr-driver-privacy/

[64] Quintin, C. and Maas, D. License plate readers exposed! How public safety agencies responded to major vulnerabilities in vehicle surveillance tech. *Electronic Frontier Foundation* (2015). https://www.eff.org/deeplinks/2015/10/license-plate-readers-exposed-how-public-safety-agencies-responded-massive

[65] Whittaker, Z. Police license plate readers are still exposed on the internet, *TechCrunch* (2019). https://techcrunch.com/2019/01/22/police-alpr-license-plate-readers-accessible-internet/

[66] Holmes, F. Over-the-air updates moving from 'nice to have' to 'vital'. *Automotive World* (2018). https://www.automotiveworld.com/articles/over-the-air-updates-moving-from-nice-to-have-to-vital/

[67] Protecting the fleet … and the car business, *KPMG* (2017). https://advisory.kpmg.us/content/dam/advisory/en/pdfs/protecting-the-fleet-web1.pdf

[68] Greenberg, A. After Jeep hack, Chrysler recalls 1.4M vehicles for bug fix. *Wired* (2015). https://www.wired.com/2015/07/jeep-hack-chrysler-recalls-1-4m-vehicles-bug-fix/

[69] Valasek, C. and Miller, C. Adventures in automotive networks and control units. *IOActive* (2014). https://ioactive.com/pdfs/IOActive_Adventures_in_Automotive_Networks_and_Control_Units.pdf

[70] Keen Security Lab of Tencent, New car hacking research: 2017, remote attack Tesla Motors again (2017).

https://keenlab.tencent.com/en/2017/07/27/New-Car-Hacking-Research-2017-Remote-Attack-Tesla-Motors-Again/

[71] See, e.g., Shifting gears in cybersecurity for connected cars, *McKinsey & Company* (2017). https://www.mckinsey.com/~/media/mckinsey/industries/automotive and assembly/our insights/shifting gears in cybersecurity for connected cars/shifting-gears-in-cyber-security-for-connected-cars.ashx ; and Kill Switch: Why connected cars can be killing machines and how to turn them off. *Consumer Watchdog* (2019). https://www.consumerwatchdog.org/sites/default/files/2019-07/KILL%20SWITCH%20%207-29-19.pdf

[72] See the Auto-ISAC website (accessed October 3, 2019). https://www.automotiveisac.com/

[73] Weimerskirch, A and Dominic, D. Assessing risk: Identifying and analyzing cybersecurity threats to automated vehicles. *University of Michigan* (2018). https://mcity.umich.edu/wp-content/uploads/2017/12/Mcity-white-paper_cybersecurity.pdf

[74] Congressional Research Service. Data protection law: An overview. *R45631* (2019). https://crsreports.congress.gov/product/pdf/R/R45631

[75] Congressional Research Service. What role might the federal government play in law enforcement reform? *IF10572* (2018). https://fas.org/sgp/crs/misc/IF10572.pdf